

# Baseball Metrics

---

Midterm Presentation

Team J

# CONTENTS

- **INTRODUCTION**
  - Objectives
  - Service Features
  - Role & Milestone
- **DESIGN**
- **FRONTEND**
  - Pages UX/UI
- **BACKEND**
  - Analysis Logic
  - Game Logic

# Objectives

AS-IS 😞

## Problems

- ✓ Frequently exposed to baseball stats but not clear
- ✓ Want a deeper understanding of stats to enjoy watching baseball
- ✓ Highly interested in evaluating players or predicting the outcome of baseball games

Solution

## Web Service

- ✓ Simple baseball line-up simulation game:
- ✓ Providing explanation of the indicators together with the records
- ✓ Visualization of information comparing the competence of the players

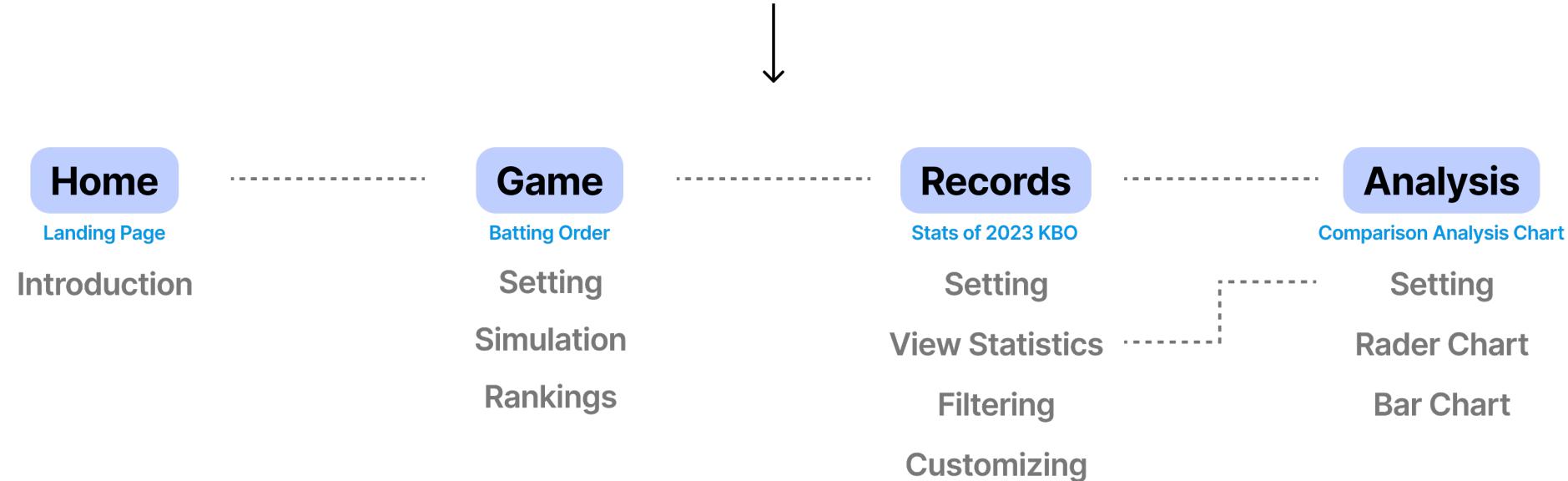
TO-BE 😊

## Goals

- ✓ Experience predicting the outcome of the game with real players
- ✓ Understand the indicators used in the broadcasting or the recording of the player
- ✓ Evaluate the players more objectively

# Service Features

## BASEBALL METRICS



# Role & Milestone : Progress

Status: In progress, Done + Add filter

Aa Task name	Status	Assignee	Due	Field
Background Research	Done	채원 고 창우 신현창	October 6, 2023	
User Research	Done	채원 고	September 22, 2023	
System Structure Design	Done	채원 고 창우 신현창	September 22, 2023	Design
Home Page Wireframe	Done	채원 고	October 13, 2023	Design
Game Page Wireframe	Done	창우	October 13, 2023	Design
Record Page Wireframe	Done	채원 고	October 13, 2023	Design
Analysis Page Wireframe	Done	채원 고	October 25, 2023	Design
Applying Design System to Wireframes	Done	채원 고	November 9, 2023	Design
Game Logic	In progress	신현창 융합생명공학과/	November 10, 2023	Backend
Determine normalization methods and constants at each stage	In progress	창우	November 10, 2023	Backend
Home Page Implementation	In progress	창우 채원 고	November 16, 2023	Frontend
Crawling the data	In progress	창우 신현창 융합생명공학	November 16, 2023	Backend
Run the code and determine constants	In progress	창우 신현창 융합생명공학	November 16, 2023	Backend
Write a short description of metrics	In progress	창우 신현창 융합생명공학	November 23, 2023	Backend

# Role & Milestone : To-do

Aa Task name	Due
▶ Complete the game function	November 17, 2023
▶ Record Page Implementation	November 23, 2023
▶ Game Page Implementation	November 23, 2023
▶ Analysis Page Implementation	November 30, 2023
▶ Pages Routing	November 30, 2023
▶ Implement ability evaluation	November 30, 2023
▶ Game Hint Logic	November 23, 2023
▶ server implementation	December 7, 2023
▶ Front-Back Connection	December 7, 2023
▶ Testing	December 11, 2023

신현창 융합생명공학과/ 5

창우 7

채원고 5

Not started 10 ... +

Complete the game function Backend

Game Hint Logic Backend

server implementation Backend

Front-Back Connection

Testing

+ New

Complete the game function Backend

Game Hint Logic Backend

Game Page Implementation Frontend

Implement ability evaluation Backend

server implementation Backend

Front-Back Connection

Testing

+ New

Record Page Implementation Frontend

Pages Routing Frontend

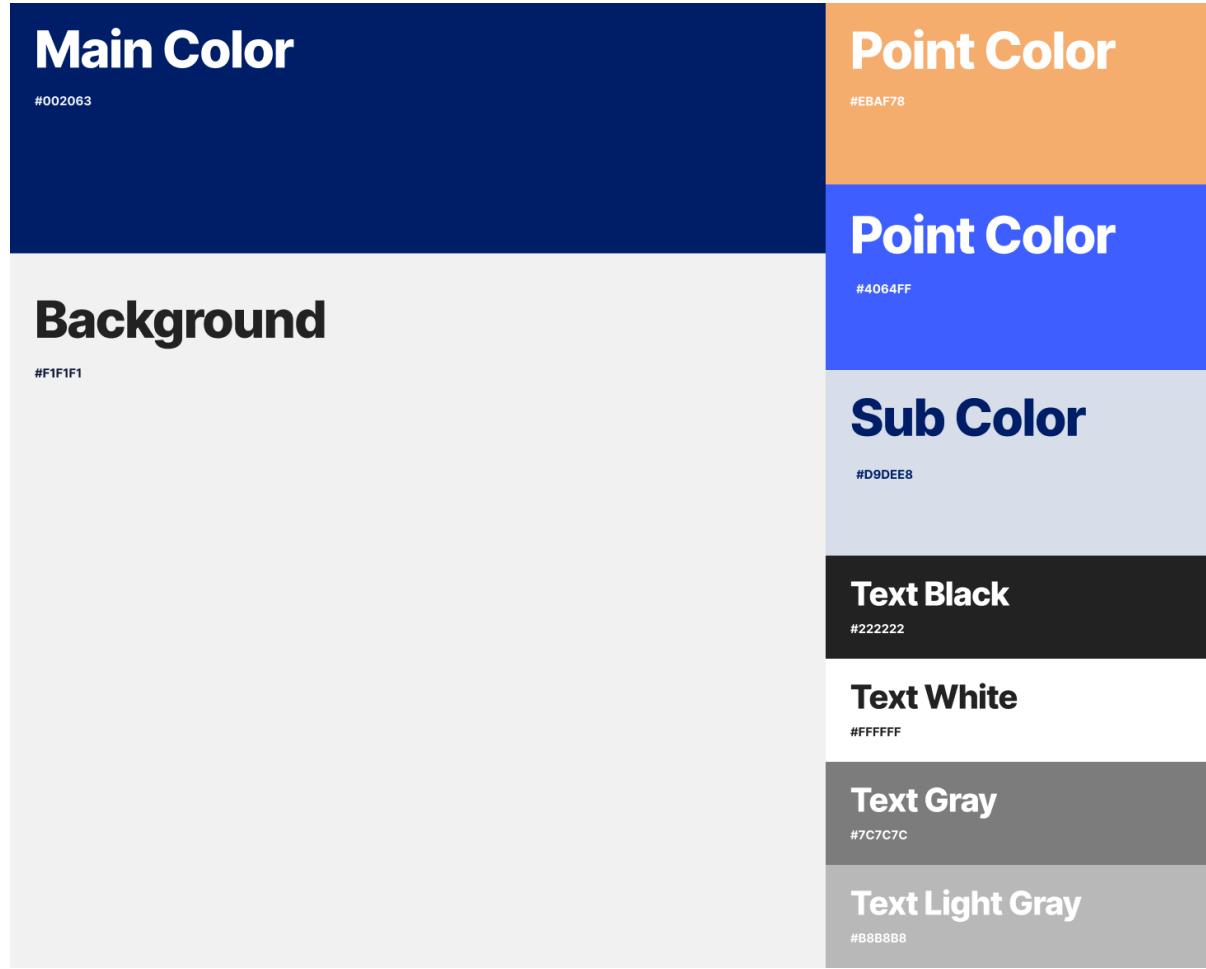
Analysis Page Implementation Frontend

Front-Back Connection

Testing

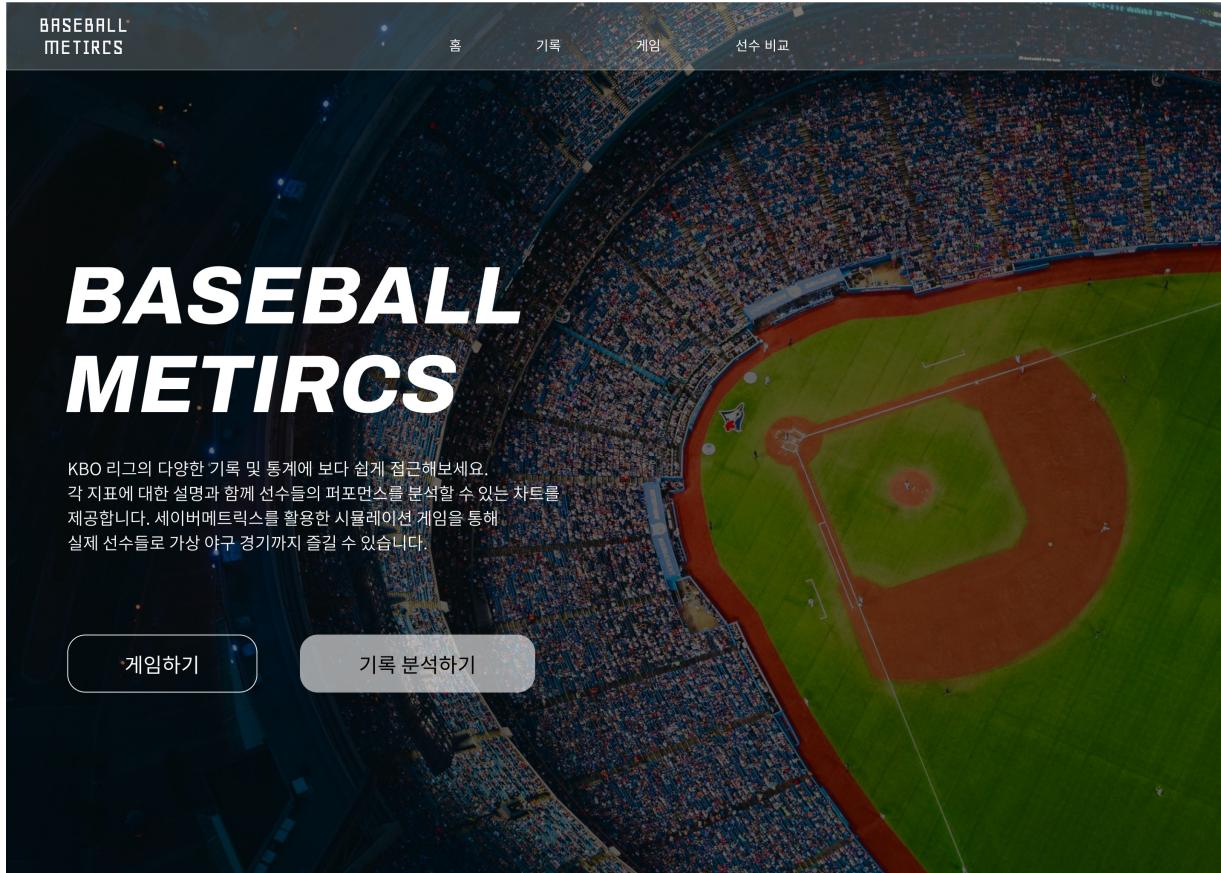
+ New

# Design



 SSG랜더스 - main #C20024	 SSG랜더스 - sub #EDB3BD
 키움하이어로즈 - main #8A0027	 키움하이어로즈 - sub #DCB3BE
 LG트윈스 - main #A50034	 LG트윈스 - sub #E4B3C2
 KT위즈 - main #EC0A0B	 KT위즈 - sub #F9B6B6
 KIA타이거즈 - main #EA0029	 KIA타이거즈 - sub #F9B3BF
 NC다이노스 - main #071D3D	 NC다이노스 - sub #C7A079
 삼성라이온즈 - main #3277C4	 삼성라이온즈 - sub #EAB821
 롯데자이언츠 - main #041E42	 롯데자이언츠 - sub #D00F31
 두산베어스 - main #131230	 두산베어스 - sub #ED1C24
 한화이글스 - main #FF6600	 한화이글스 - sub #FFD1B3

# Frontend Home Page



- Introduction to the website
- Route to two main features: Game and Analysis

# Frontend Game Page : Pitcher Selection

The screenshot shows a game page interface. At the top, there is a navigation bar with tabs: 허 (Home), 게임 (Game), 기록 (Record), and 선수 비교 (Player Comparison). Below the navigation bar is a section titled "상대 투수를 선택해주세요" (Select the opposing team's pitcher). This section contains a table of pitcher statistics for the Daegu Samsung Lions team. The table includes columns for 순위 (Rank), 선수명 (Player Name), 팀명 (Team), W-L, BB, SO, GDP, SLG, OBP, BA/RISP, 대타타율 (Batter's batting average), GO/AO, and BB/K. The table lists 15 players, with the 4th player, Kim Tae-ho, highlighted in blue. Below the table is a section titled "구단을 선택해주세요" (Select the league) featuring logos for various Korean baseball teams: LG Twins, KT Wiz, SSG Landers, NC Dinos, Doosan Bears, KIA Tigers, Lotte Giants, Hanwha Eagles, Samsung Lions, Kiwoom Heroes, and ALL. At the bottom left, there is a section titled "타자를 선택해주세요" (Select the batter).

순위	선수명	팀명	W-L	BB	SO	GDP	SLG	OBP	BA/RISP	대타타율	GO/AO	BB/K
1	양의지	두산	0.462	0	1	2	0	0	0	0	0	0
2	홍성호	두산	0.400	4	3	0	0	0	0	0	0	0
3	정수빈	두산	0.400	1	2	0	0	0	0	0	0	0
4	김재호	두산	0.385	0	3	0	0	0	0	0	0	0
5	양식환	두산	0.333	2	2	0	0	0	0	0	0	0
6	허경민	두산	0.333	2	2	1	0	0	0	0	0	0
7	박유연	두산	0.308	2	2	0	0	0	0	0	0	0
8	강승호	두산	0.250	4	1	0	0	0	0	0	0	0
9	김인태	두산	0.222	4	2	0	0	0	0	0	0	0
10	로하스	두산	0.167	0	2	0	0	0	0	0	0	0
11	이유찬	두산	0.167	2	1	0	0	0	0	0	0	0
12	양찬열	두산	0.167	1	1	0	0	0	0	0	0	0
13	전민재	두산	0.111	3	3	1	0	0	0	0	0	0
14	송승환	두산	0.091	2	3	0	0	0	0	0	0	0
15	박준영	두산	0.091	0	4	0	0	0	0	0	0	0

- Select the opposing team's pitcher (one)
  - Metrics of pitcher we use here:  
Win-lose, innings per game (total innings/entrance), earned run, hit, home run, strikeout, walk, ERA, WHIP, pitch type
  - Mouse Hover on Indicators: One-line description of the indicator “*It is calculated based on (recorded values) and has (relation) to (athlete's performance).*”
- Limitations: Reliability and Consistency**

- Stats Header: fixed
- Stats Number: Scrollable

# Frontend Game Page : Batter Selection

구단을 선택해주세요

LG TWINS KT KIA SSG LANDERS NC BEARS TIGERS DOLPHINS 롯데 한화 Lions 삼성 키움 ALL

타자를 선택해주세요

순위	선수명	팀명	Avg	BB	SO	GDP	SLG	OBP	BA/RISP	phLI	GO/AO	BB/K
1	양의지	두산	0.462	0	1	2	0	0	0	0	0	0
2	홍성호	두산	0.400	4	3	0	0	0	0	0	0	0
3	정수빈	두산	0.400	1	2	0	0	0	0	0	0	0
4	김제호	두산	0.385	0	3	0	0	0	0	0	0	0
5	양식환	두산	0.333	2	2	0	0	0	0	0	0	0
6	허경민	두산	0.333	2	2	1	0	0	0	0	0	0
7	박유연	두산	0.308	2	2	0	0	0	0	0	0	0
8	강승호	두산	0.250	4	1	0	0	0	0	0	0	0
9	김인태	두산	0.222	4	2	0	0	0	0	0	0	0
10	로하스	두산	0.167	0	2	0	0	0	0	0	0	0
11	이유찬	두산	0.167	2	1	0	0	0	0	0	0	0
12	양찬열	두산	0.167	1	1	0	0	0	0	0	0	0
13	전민재	두산	0.111	3	3	1	0	0	0	0	0	0
14	송승환	두산	0.091	2	3	0	0	0	0	0	0	0
15	박준영	두산	0.091	0	4	0	0	0	0	0	0	0

선택한 타자의 태순을 정해주세요 ⓘ 태순은 어떻게 정하나요?

1 두산 2 두산 3 두산 4 두산 5 두산 6 두산 7 두산 8 두산 9 두산

- Select your team if any
- Select 9-15 batters
- Click to Select
- Click again to deselect
- Selected ones appear below in the order you click
- Metrics of batters we use here:  
AVG(타율), BB(볼넷), SO(삼진), GDP(병살타),  
SLG(장타율), OBP(출루율), BA/RISP(득점권타율),  
phLI(대타LI), GO/AO(땅볼/뜬공), BB/K(볼넷/삼진)
- Mouse Hover on Indicators: One-line description of the indicator  
*“It is calculated based on (recorded values) and has (relation) to (athlete’s performance).”*
- Stats Header: fixed
- Stats Number: Scrollable

# Frontend Game Page : Ordering Batters

12	양찬열	두산	0.167	1	1	0	0	0	0	0	0	0
13	전민재	두산	0.111	3	3	1	0	0	0	0	0	0
14	송승환	두산	0.091	2	3	0	0	0	0	0	0	0
15	박준영	두산	0.091	0	4	0	0	0	0	0	0	0

선택한 타자의 타순을 정해주세요 (?) 타순은 어떻게 정하나요?

선택한 타자의 타순을 정해주세요 (?) 타순은 어떻게 정하나요?

1	두산											
5	두산	5	6	7	8	9	10	11	12	13	14	15
1	양의지											
10	양의지											
11	대타											
12	양의지											
13	페디											
14	페디											
15	페디											

시뮬레이션 시작하기

- Drag to determine the batting orders of the selected batters
- Hint: Provide information about each batting order

# Frontend Game Page : Running Simulation

The screenshot shows a baseball game simulation interface with the following elements:

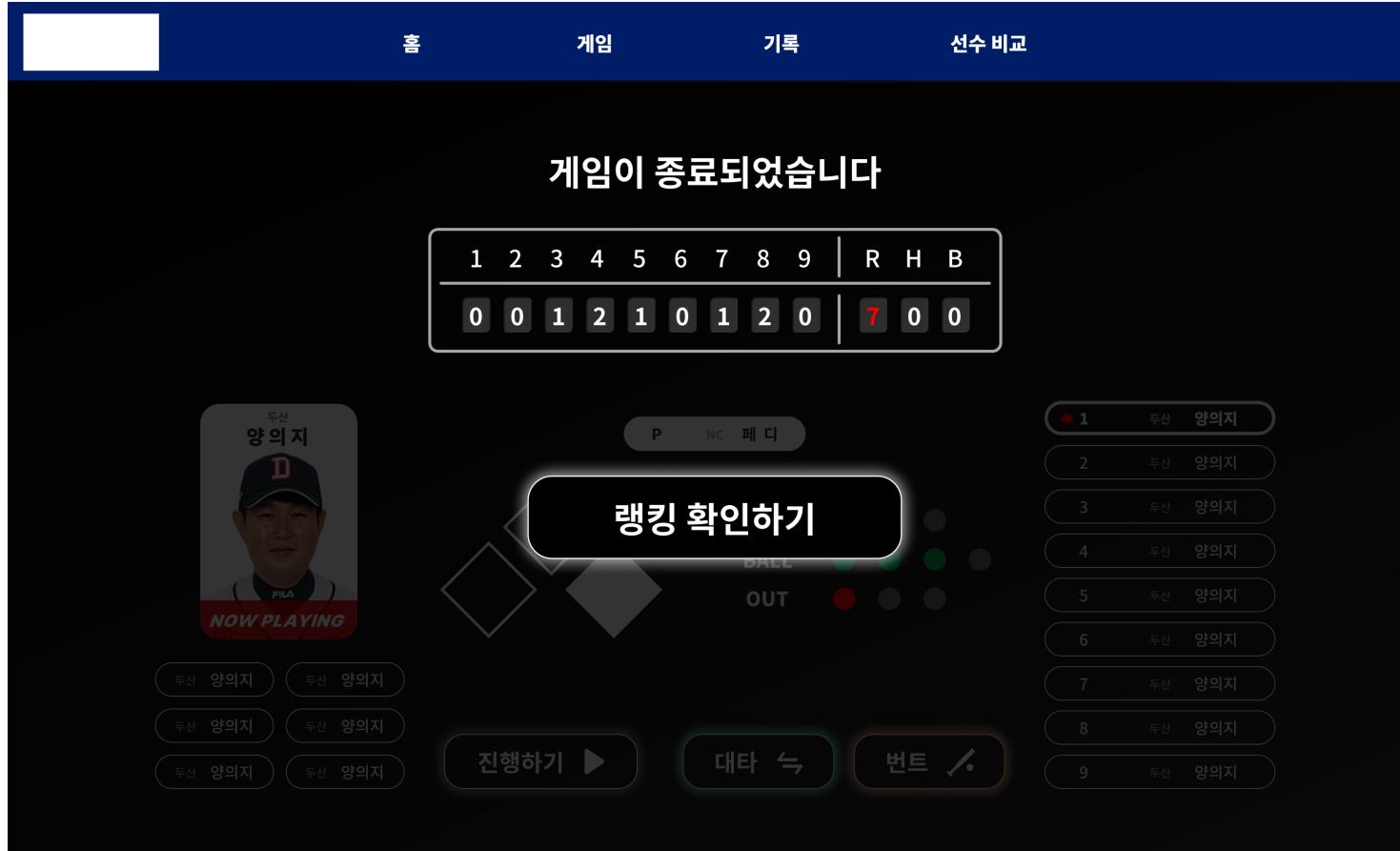
- Header:** 홍 (Home), 게임 (Game), 기록 (Record), 선수 비교 (Player Comparison).
- Middle Top:** 게임이 진행 중입니다 (Game is in progress).
- Middle Center:** Scoreboard table:

1	2	3	4	5	6	7	8	9	R	H	B
0	0	1									
- Middle Left:** Current player: 양의지 (Yang I-ji) from 두산 (Doosan), status: NOW PLAYING.
- Middle Center:** On-base situation: S, B, O counting. A tip message: Tip! 어쩌구 저쩌구 한 상황입니다. 어쩌구 저쩌구를 고려하세요.
- Middle Right:** Batting orders: current player highlighted (red dot).

1	두산	양의지
2	두산	양의지
3	두산	양의지
4	두산	양의지
5	두산	양의지
6	두산	양의지
7	두산	양의지
8	두산	양의지
9	두산	양의지
- Bottom Left:** Pinch-hitters: Abled when you click "pinch-hitter" btn.
- Bottom Center:** Action buttons: 진행하기 (Proceed), 대타 (Pinch-hitter), 번트 (Bunt).
- Bottom Right:** After each at-bat, decide how to proceed the game: Refer to the tip!

\*at-bat(타석): until one batter completes the batting action with an out or on base

# Frontend Game Page : End of Simulation



- After 9 innings, game ends
- Result of your game is shown
- Check out for your ranking

# Frontend Game Page : User Ranking

The screenshot shows a game ranking page with the following data:

Rank	User ID	Score	Statistics
5	USER ID-5	15점	1 2 0 2 1 2 0 0 3 1 1 2
1	USER ID-1	15점	1 2 0 2 1 2 0 0 3 1 1 2
2	USER ID-2	15점	1 2 0 2 1 2 0 0 3 1 1 2
3	USER ID-3	15점	1 2 0 2 1 2 0 0 3 1 1 2
4	USER ID-4	15점	1 2 0 2 1 2 0 0 3 1 1 2
5	USER ID-5	15점	1 2 0 2 1 2 0 0 3 1 1 2
6	USER ID-5	15점	1 2 0 2 1 2 0 0 3 1 1 2

A note at the top states: "투수 페디와 플레이한 유저들의 득점 순위입니다."

- You can compare your scores with others who played against the same pitcher you selected
- If the total score is the same, the ranking differs by the number of homeruns and hits

# Frontend Record Page

The screenshot shows a web application interface for a baseball record page. At the top, there are four small cards representing game results from October 27, 2023, between LANDERS and BEARS, each showing a score of 5-3. Below these are three tabs: '기록' (Record), '선수 비교' (Player Comparison), and '선수 통계' (Player Statistics). The '기록' tab is selected.

Under the '기록' tab, there is a table of player statistics. The columns include 순위 (Rank), 선수명 (Player Name), 팀명 (Team), W-L, BB, SO, GDP, SLG, OBP, BA/RISP, 대타타율 (Bunting Rate), GO/AO, and BB/K. The data shows 14 players from the LANDERS team, all named 양의지 (Yang Eui-ji), with a batting average of 0.462 across various categories.

A tooltip is visible over the '한글 단어' (Korean Word) column, explaining that it uses '@@' to calculate metrics like BA/RISP and provides a link to learn more.

순위	선수명	팀명	W-L	BB	SO	GDP	SLG	OBP	BA/RISP	대타타율	GO/AO	BB/K
1	양의지	두산	0.462	0	1	2	0	0	0	0	0	0
2	양의지	두산	0.462	0	1	2	0	0	0	0	0	0
3	양의지	두산	0.462	0	1	2	0	0	0	0	0	0
4	양의지	두산	0.462	0	1	2	0	0	0	0	0	0
5	양의지	두산	0.462	0	1	2	0	0	0	0	0	0
6	양의지	두산	0.462	0	1	2	0	0	0	0	0	0
7	양의지	두산	0.462	0	1	2	0	0	0	0	0	0
8	양의지	두산	0.462	0	1	2	0	0	0	0	0	0
9	양의지	두산	0.462	0	1	2	0	0	0	0	0	0
10	양의지	두산	0.462	0	1	2	0	0	0	0	0	0
11	양의지	두산	0.462	0	1	2	0	0	0	0	0	0
12	양의지	두산	0.462	0	1	2	0	0	0	0	0	0
13	양의지	두산	0.462	0	1	2	0	0	0	0	0	0
14	양의지	두산	0.462	0	1	2	0	0	0	0	0	0

- Set whose stats to view: team, batter or pitcher
- Set which stats to view: select a category
- Mouse Hover on Indicators: One-line description of the indicator
- Live match broadcasting if any
- You can link to analysis page if you need the player's comparison analysis

# Frontend Record Page : Filtering Index

The screenshot shows a baseball record page with the following sections:

- Game Results:** Four boxes showing games from October 27, 2023, between LANDERS and BEARS. All games ended in a 5-3 victory for LANDERS.
- Filtering Options:** Buttons for "전체 팀" (All Team), "타자" (Batter), and "투수" (Pitcher). A "스탯 커스텀" (Stat Custom) button is also present.
- Player Statistics Table:** A table with 14 rows of data. The columns include 순위 (Rank), 선수명 (Player Name), 팀명 (Team), W-L, BB, SO, GDP, SLG, OBP, BA/RISP, and 대 (Opponent). The data shows various players from the Doosan Bears (두산) team with similar statistics across all games.
- Filter Overlay:** A modal window titled "필터" (Filter) allows setting ranges for W-L, BB, and SO counts.

순위	선수명	팀명	W-L	BB	SO	GDP	SLG	OBP	BA/RISP	대
1	양의지	두산	0.462	0	1	2	0	0	0	
2	양의지	두산	0.462	0	1	2	0	0	0	
3	양의지	두산	0.462	0	1	2	0	0	0	
4	양의지	두산	0.462	0	1	2	0	0	0	
5	양의지	두산	0.462	0	1	2	0	0	0	
6	양의지	두산	0.462	0	1	2	0	0	0	
7	양의지	두산	0.462	0	1	2	0	0	0	
8	양의지	두산	0.462	0	1	2	0	0	0	
9	양의지	두산	0.462	0	1	2	0	0	0	
10	양의지	두산	0.462	0	1	2	0	0	0	
11	양의지	두산	0.462	0	1	2	0	0	0	
12	양의지	두산	0.462	0	1	2	0	0	0	
13	양의지	두산	0.462	0	1	2	0	0	0	
14	양의지	두산	0.462	0	1	2	0	0	0	

Filter specific range of each stats to view

# Frontend Record Page : Customizing Index

The screenshot shows a baseball record page with a navigation bar at the top: 홈 (Home), 게임 (Game), 기록 (Record), and 선수 비교 (Player Comparison). Below the navigation bar, there are four game cards for October 27, 2023, between LANDERS and BEARS, all resulting in a 5-3 win for LANDERS.

A modal window titled "보고싶은 스탯을 선택해주세요" (Select the stats you want to see) is displayed. It lists various statistics categories and their checkboxes:

- 기본**: W-L (unchecked), BB (checked), SO (checked), GDP (unchecked), SLG (unchecked), OBP (checked), BA/RISP (unchecked), BB/K (unchecked)
- 확장**: W-L (unchecked), BB (unchecked), SO (unchecked), GDP (unchecked), SLG (checked), OBP (checked), BA/RISP (unchecked), BB/K (unchecked)
- 클러치**: W-L (checked), BB (unchecked), SO (unchecked), GDP (checked), SLG (checked), OBP (unchecked), BA/RISP (unchecked), BB/K (checked)
- 타구1**: W-L (checked), BB (checked), SO (checked), GDP (checked), SLG (unchecked), OBP (checked), BA/RISP (unchecked), BB/K (unchecked)

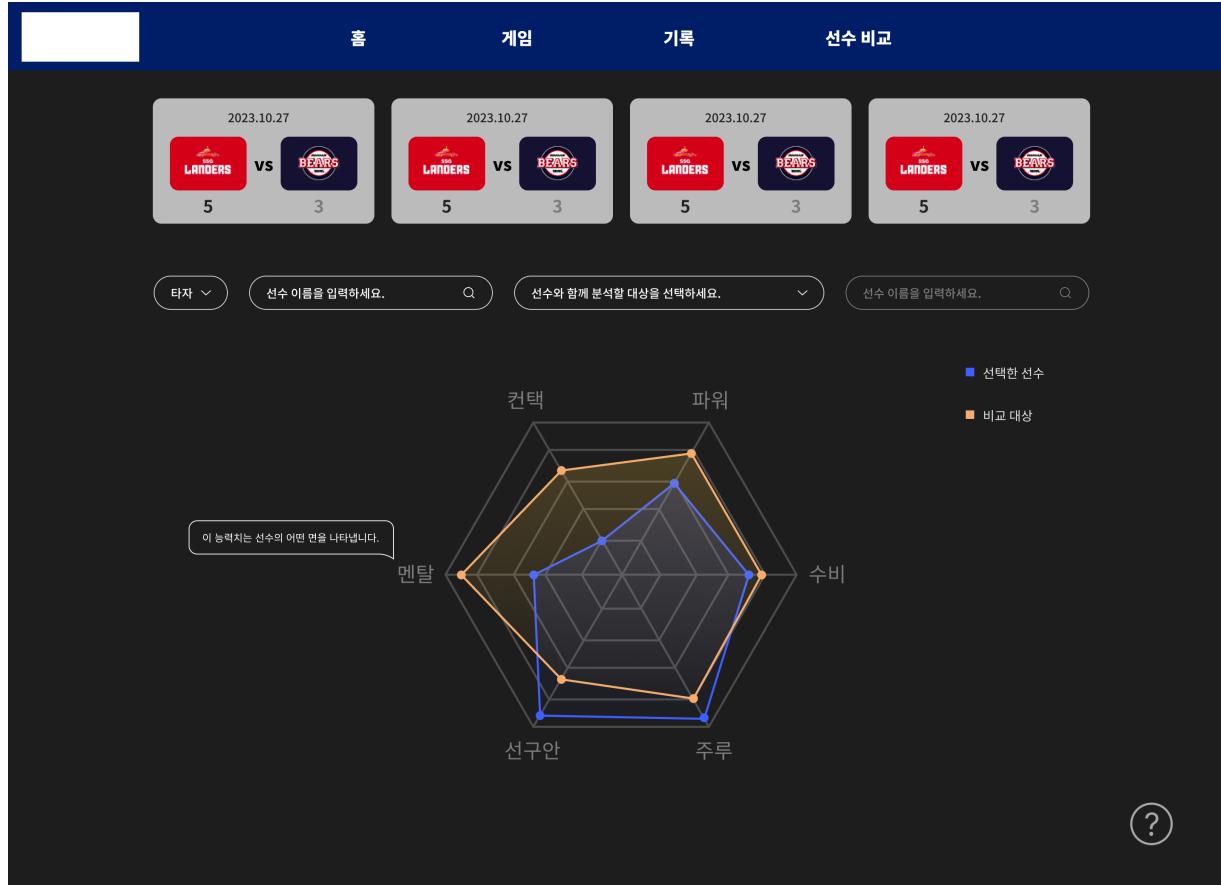
Below the modal, a tooltip provides information about the search function: "한글 단어 @@@를 통해 계산한다. 이를 통해 &&를 파악할 수 있다." (Korean words @@@ are calculated. You can identify && through them.)

The main table below the modal displays player statistics for two players:

번호	선수명	팀	W-L	BB	SO	GDP	SLG	OBP	BA/RISP	BB/K
13	양의지	두산	0.462	0	1	2	0	0	0	0
14	양의지	두산	0.462	0	1	2	0	0	0	0

- Customize which stats to view
- The index of the table is composed of the selected items

# Frontend Analysis Page : Performance Analysis



- Select a player that you want to analyze
- Select a comparison target to be analyzed with the player: the average of one's team, kbo league average, another player
- Performance Analysis:
  - Presented with a radar chart
  - 6 attributes for batters
  - 5 attributes for pitchers
- Mouse hover on each attribute: display meaning of the attribute
- Show how each attribute is calculated using which stats

# Frontend Analysis Page : Metrics Comparison



The stat values of the two targets are compared and shown as a bar chart

# 6 Abilities of Batter

## **Power: How far the batter can send the ball**

Homerun, SLG(how many bases a batter reach per at-bat), Spd (ability of speed)

## **Contact: How well the batter hits the ball and precisely**

Contact%, cut after 2S% (Cut: contact the ball but foul)

## **Batting Eye: How well the batter can distinguish between a ball and a strike**

Batting eye after 2S%, IsoD(pure on-base percentage = on-base% - AVG%)

## **Mentality: How well the batter is in important situations**

Clutch(ability at important situation), WPA(win probability added)

## **Speed: Running speed of the batter and the sense of base running**

Spd, Steal(1B→2B while pitch), F23(probability of 2B→3B at fly out), H13(probability of 1B→3B at single hits)

## **Defense: The batter's ability to defend.**

Defense RAA(ability of defense)

The indicators below description that it is difficult to grasp the meaning at a glance and the numerical values are not easily interpretable. Therefore, we aim to convert them into representative metrics for intuitive comparison and evaluation.

# 5 Abilities of Pitcher:

**Ball Power:** Pitcher's ability to disrupt the batter using various trajectories and techniques

K/9, LSO%, SwStr%, Contact\_IZ%, Swing\_OZ%

**Ball Speed:** How fast the pitcher can throw a ball.

Fastball Speed

**Ball Control:** Pitcher's ability to pitch exactly where he wants to be

BB/9, LSO%

**Mentality:** Mental ability to maintain performance with stable mental strength even in difficult situations

LOB%, WHIP, Double play rate

**Strength:** How many innings a pitcher can take and how many pitches he can pitch

**Starter:** IP/G(Innings per game) P/G(pitches per game)

**Relief:** G, P/IP

The pitcher's 5 abilities are also converting difficult-to-recognize indicators into intuitive abilities.

# How can we convert indicators to 11 abilities?

First, normalization indicators. Use **min-max normalization**

(Limiting the minimum number of data samples for a player can remove outliers.

Second, for several indicators, use **weighted average**.

Lastly, **convert the result to 20~80** (intuitive score)

**Power: Homerun, SLG**(how many bases a batter can reach per at-bat), **Spd**(Speed ability)

$$pHR = \frac{HR - \min(HR)}{\max(HR) - \min(HR)} \quad pSLG = \frac{SLG - \min(SLG)}{\max(SLG) - \min(SLG)} \quad pSpd = \frac{Spd - \min(Spd)}{\max(Spd) - \min(Spd)}$$

$$Power = 20 + 60 \cdot \frac{w_1 pHR + w_2 pSLG + w_3 pSpd}{w_1 + w_2 + w_3}$$

**For example,**

HR is the most important value for power, so let  $w_1=6$ , SLG means how many bases a batter can reach per at-bat and it's also important, so let  $w_2=5$ . But fast batter can get 2B by legs, so let  $w_3=-1$ .

$$Power = 20 + 60 \cdot \frac{5pHR + 6pSLG + (-1)pSpd}{10}$$

# How modified?

## Logistic Function

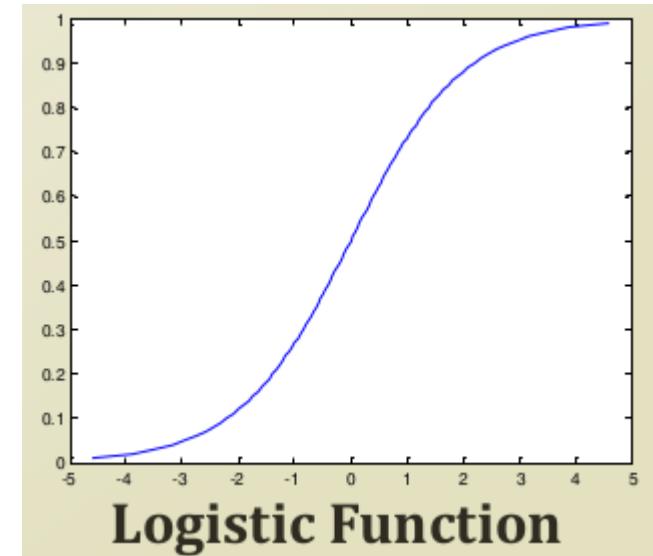
Let  $S$  = season stat,  $H$  = head-to-head stat. we use logistic function that

$$\text{Modified stat} = S - \left( \frac{S - H}{1 + e^{-k(n-8)}} \right)$$

*n is sample size of H*  
*k is a constant*

## What we care about:

- Small sample will have little impact. Large sample will have big impact.
- Function must be bounded.
- Size of head-to-head stat is from 0 to 16 in the ordinary case (1season)
- Therefore, we use sigmoid(logistic) function and let inflection point to n=8



## Limitation

- If no head-to-head stat?

```
def sigmoid(average, vsAverage, num):  
    return average - (average-vsAverage)/(1+m.exp(-0.5*(num-8)))
```

```
playerK = sigmoid(playerK, playerK + 50 , 15)  
playerBB = sigmoid(playerBB, playerBB - 50 , 15)
```

# How modified?

## Weighted Average

Let B = batter's infield% P = pitcher's infield%

$$\text{Modified infield\%} = \frac{mB+nP}{m+n}$$

(m,n) = (2,3). can be changed

## What we care about:

- The pitcher's ability to induce ground balls has a greater impact on the game
- Nevertheless, the ability of the batter to produce fly balls is also important

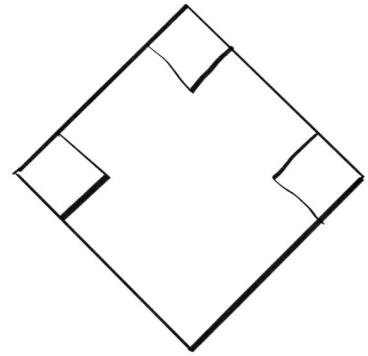
## Limitation

- Pitchers' ground ball induction is closely related to pitches such as changeups and sinkers, but it is difficult to consider pitches.

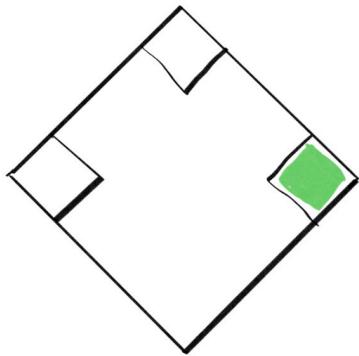
# Base management

Base :  $000_2 \sim 111_2$ ,  $2^2 \cdot 3B + 2^1 \cdot 2B + 2^0 \cdot 1B$

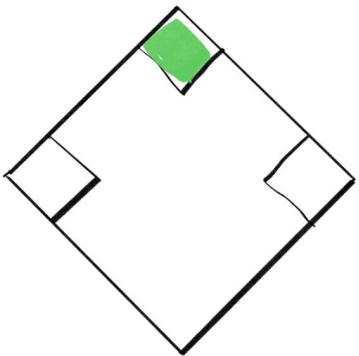
(where runner on the base: 1, no runner on the base: 0)



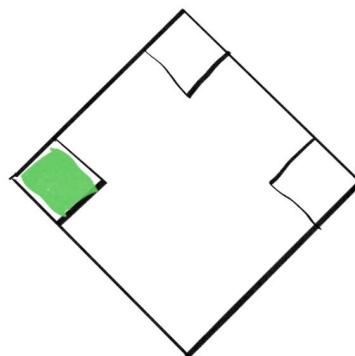
$000$   
0



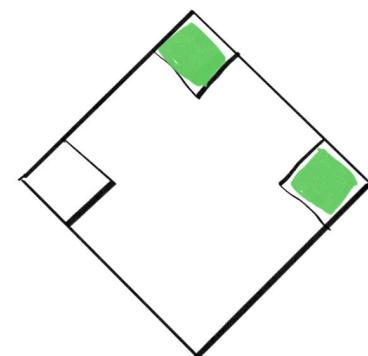
$001$   
1



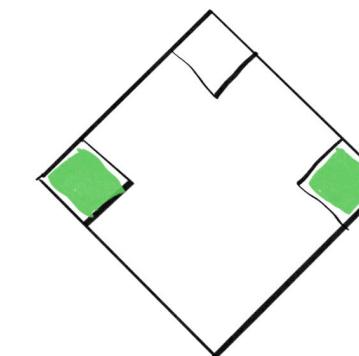
$010$   
2



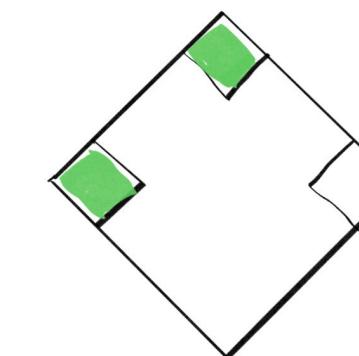
$100$   
4



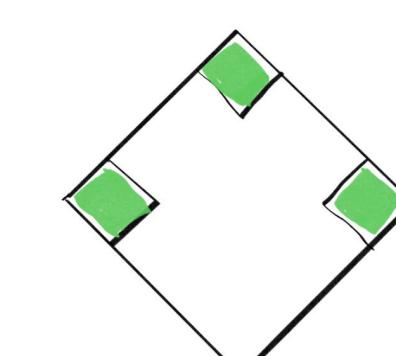
$011$   
3



$101$   
5

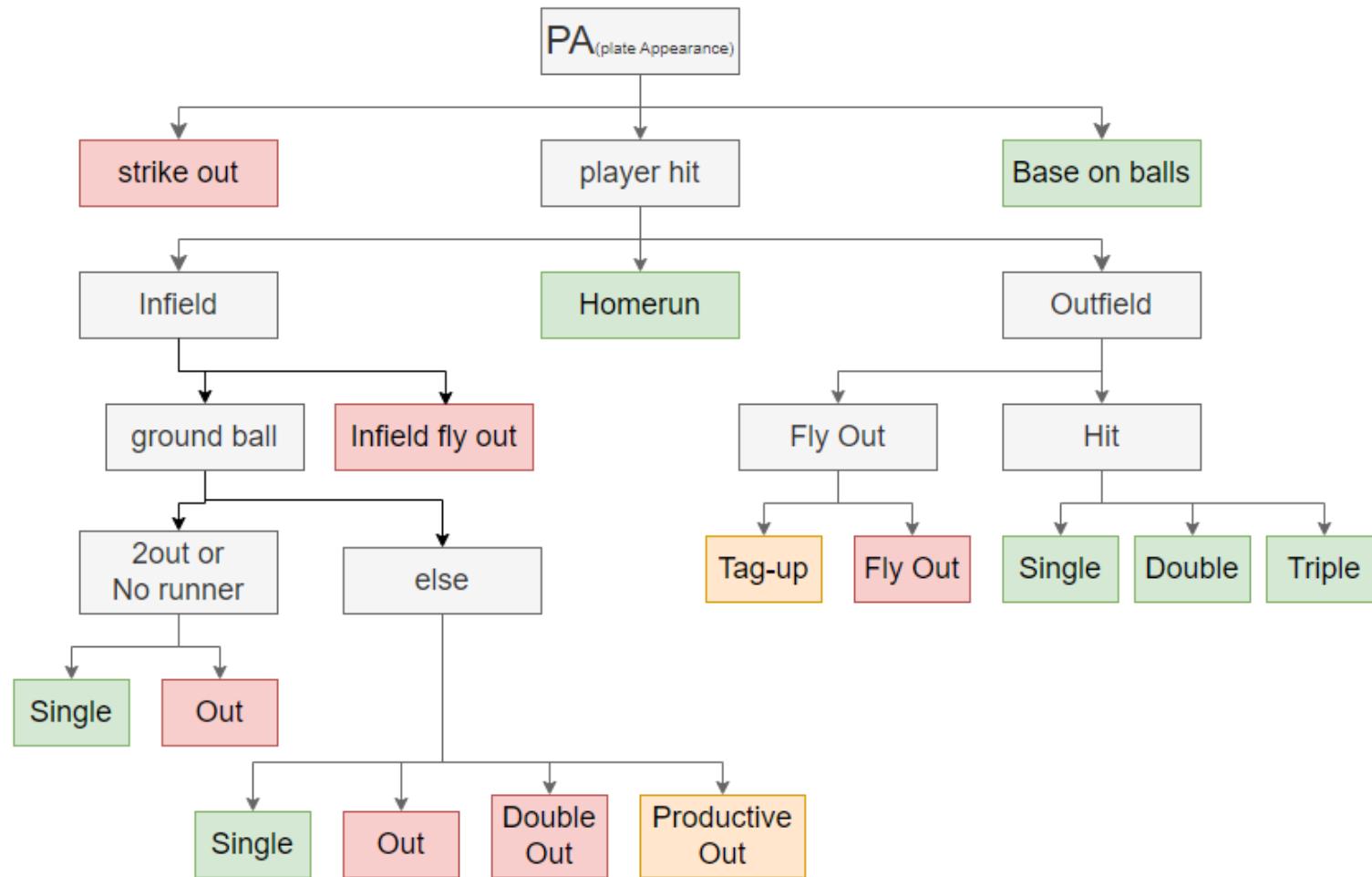


$110$   
6

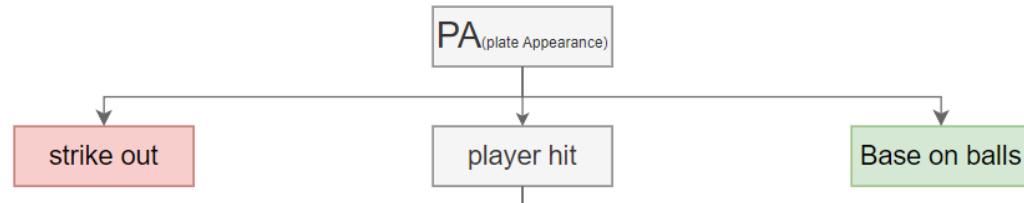


$111$   
7

# Game Logic



# K, BB, Player Hit(Contact)



## [Used data/indicators]

Batter's season K%, BB%

Head-to-head K%, BB%

(K%: strike out per PA, BB%: base on balls per PA, PA: plate appearance)

## [Function]

For modified K%, strike out

For modified BB%, base on balls

Else, player hit (here, player hit means contact including out and safe hit)

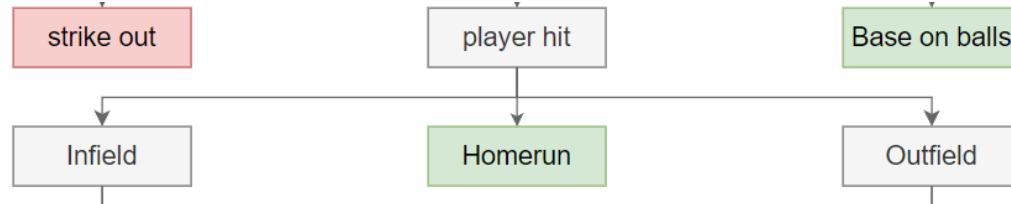
```
def hit_K_BB(name, pitcher, base, outcount):
    playerK = exp.loc[exp["이름"] == name]['K%'] * 10
    playerBB = exp.loc[exp["이름"] == name]['BB%'] * 10

    # vs 결과가 5푼 정도 차이가 날 때
    playerK = sigmoid(playerK, playerK + 50, 15)
    playerBB = sigmoid(playerBB, playerBB - 50, 15)

    playerhit = 1000 - playerK - playerBB
    -----
    selector = r.randrange(1,1001)

    if (selector <= int(playerK)) :
        outcount += 1
        return 'strikeout', base, outcount
    elif (selector <= int(playerK + playerBB)) :
        if base&1 == 1:
            base += 1
        elif base&1 == 1 and base&2 == 2:
            base += 3
        elif base&1 == 1 and base&2 == 2 and base&4 == 4:
            base += 7
        base += 1
        return 'BB', base, outcount
    else :
        return isHomerun(name,pitcher,base,outcount)
```

# HR, Infield, Outfield



## [Used data/indicators]

Season HR%, Total infield batting %

Head-to-head HR%

$$HR\% = \frac{HR}{PA \cdot InP\%}$$

PA: Plate Appearance, InP%: Inplay ball percentage

## [Function]

For modified HR%, Homerun

For modified Infield%, Infield

Else, Outfield

```
def isHomerun(name, pitcher, base, outcount):
    playerHit1 = hit1.loc[hit1["이름"] == name]
    playerHrNum = basic.loc[basic["이름"] == name]['홈런']
    playerInp = playerHit1['InP%']
    playerInp = playerInp.astype(float)
    playerPA = playerHit1['타석']
    playerHitNum = int(playerInp * playerPA / 100)
    playerHrRatio = playerHrNum/playerHitNum
    playerHrRatio *= 1000

    playerInfieldRatio = playerHit1['방향_내야%']
    playerInfieldRatio = playerInfieldRatio.astype(float)
    playerInfieldRatio *= 10

#homerun vs 결과가 5푼 정도 차이 날 때
playerHrRatio = sigmoid(playerHrRatio,playerHrRatio+50,10)

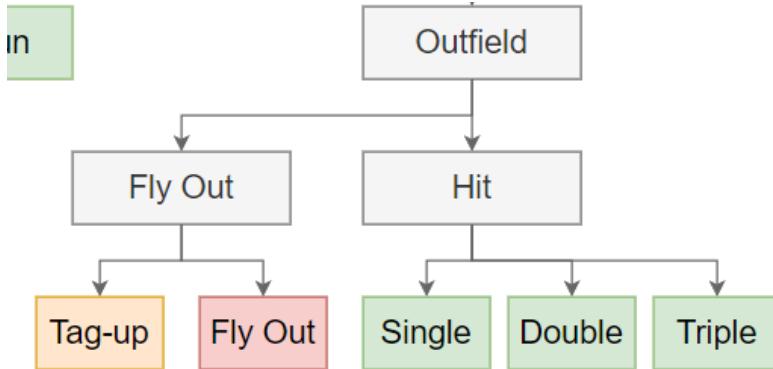
#infieldRatio 비교는 비율로 계산
playerInfieldRatio = (playerInfieldRatio*2 + (playerInfieldRatio+50)*3)/5

playerOutfieldRatio = 1000 - playerInfieldRatio.values[0] - playerHrRatio.values[0]

selector = r.randrange(1,1001)

if(selector <= int(playerHrRatio)) :
    base = base << 4
    base += 8
    return 'homerun' , base, outcount
elif(selector <= playerHrRatio.values[0] + playerInfieldRatio.values[0]):
    return infield(name,base,outcount)
else:
    return outfield(name,pitcher,base,outcount)
```

# Outfield: Fly out, Hits



[Used data/indicators]

Season Outfield AVG, AVG, Head-to-head AVG

[Modify Outfield AVG]

There are no head-to-head outfield avg. So, we compare AVG and head-to-head AVG and use the result at outfield AVG  
Calculate the value of **sigmoid(AVG,H\_AVG)-AVG** and **add it**.

[Function]

For Modified Outfield AVG%, Hit

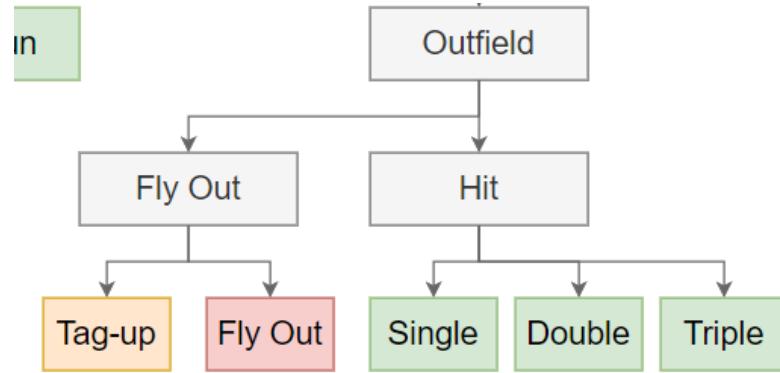
Else, Fly Out(we have to consider tag-up, next page)

```
def outfield(name,pitcher,base,outcount):
    playerOutfieldBA = float(hit1.loc[hit1["이름"] == name]['타율_외'].values[0])
    playerBA = float(basic.loc[basic["이름"] == name]['타율'].values[0])
    selector = r.randrange(1,1001)
    playerOutfieldBA *= 1000
    playerBA *= 1000

    #고민사항, 외야시 안타 비율로 안타 계산하는데 어떻게 할것인지
    # -> 일단 그냥 타율 비교하여 차이 구하고 그걸로 조정
    playerOutfieldBA += sigmoid(playerBA, playerBA+50, 15) - playerBA

    if(selector <= playerOutfieldBA):
        return outfieldHit(name,pitcher,base,outcount)
    else:
        return outfieldFO(name,base,outcount)
```

# Fly out: Tag-up and Fly out



## [Used data/indicators]

Right-handed hit rate%

## [Function]

If 2out, fly out

Right-handed hit rate% -> right-handed hit

3B runner: advanced regardless of batted ball direction

2B runner: advanced if right-handed hit

## [Limitation]

Runner's speed and Short right-handed hit are not considered.

```
def outfieldFO(name, base, outcount):
    if(outcount == 2):
        outcount += 1
        return 'out',base,outcount
    outcount += 1

    playerOutfieldCourse = float(hit2.loc[hit2["이름"] == name]['타구방향%_우'].values[0])
    playerOutfieldCourse *= 10

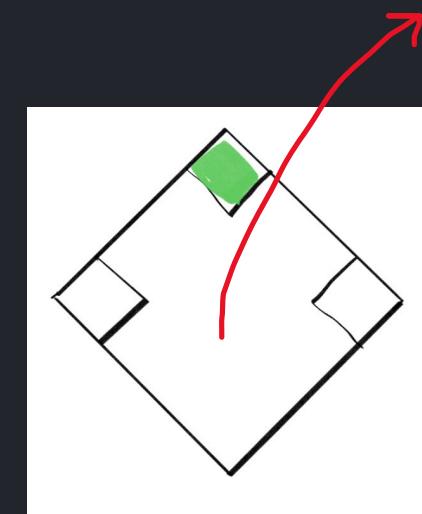
    course = 0
    selector = r.randrange(1,1001)

    if(selector <= playerOutfieldCourse):
        course = 1

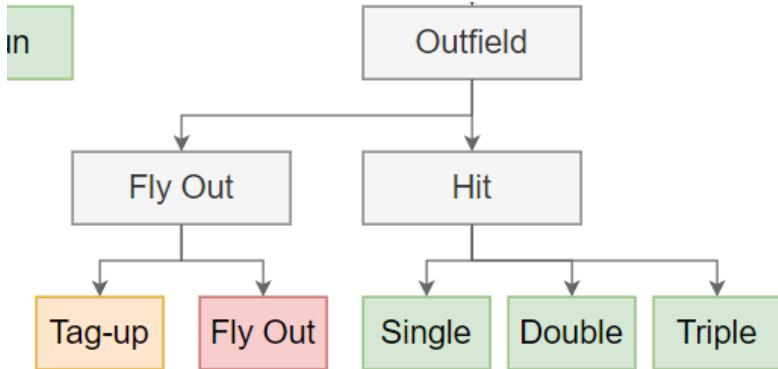
    result = 'outFO+'

    if(base & 4 == 4):
        base += 4
        result += '34'
    if((base & 2 == 2) and (course)):
        base += 2
        result += '23'

    return result, base, outcount
```



# Outfield: Single, Double, Triple



## [Used data/indicators]

Right/Left-handed hit rate%, #Hits, #Infield hits, #Double, #Triple, additional advance probability at hits%(H+%)

$\frac{(N \text{ base hits})}{(\text{Hits}) - (\text{Infield hits})}$  : calculate **N base hits probability**

## [Function]

Right/Left%=> determine direction of the hit

For N base hits probability, N base hits(single, double, triple)

Runner is decided by H+% and course(direction of the hit)

```
# course = 2 left, course = 1 right, course = 0 center

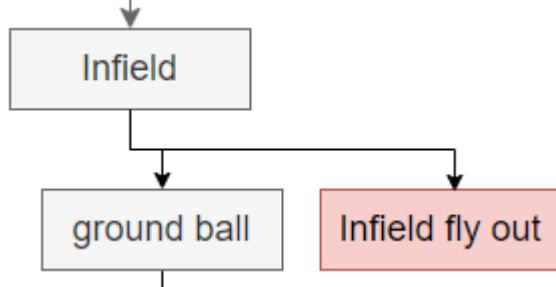
if(selector2 <= playerHitLeft) :
    course = 2
elif(selector2 <= playerHitLeft + playerHitRight) :
    course = 1
else :
    course = 0

if(selector <= int(player1BRatio)):
    if(course == 1):
        base = base << 2
        base += 1
        return '1B+', base, outcount
    else:
        base = base << 1
        base += 1
        return '1B', base, outcount

elif(selector <= int(player1BRatio + player2BRatio)):
    if(selector3 <= int(playerHitplus)):
        base = base << 3
        base += 2
        return '2B+', base, outcount
    else:
        base = base << 2
        base += 2
        return '2B', base, outcount

else:
    base = base << 3
    base += 3
    return '3B', base, outcount
```

# Infield fly and groundball



**Groundballs have a greater risk of double play or out compared to straight hits. So, we think about it separately.**

## [Used data/indicators]

#Infield fly out, Inplay rate(InP%), PA, Infield hits rate%(Infield%)

## [Function]

Calculate #Hit, #Infield hit, rate of infield fly%

#Hit = PA \* InP%, #Infield hit = (Infield% \* #Hit) / 100

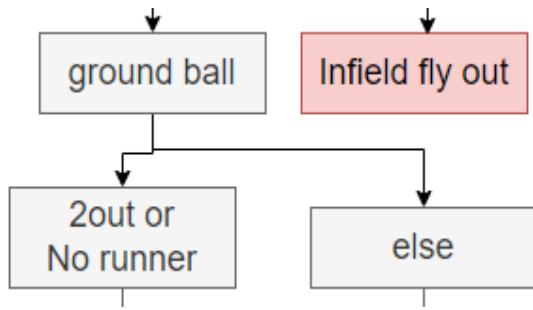
Rate of infield fly% = Infield fly out / #Infield hit

Calculate #Infield groundball = #Infield hit - #infield fly  
(Use later)

For rate of infield fly%, Infield fly  
Else, groundball

```
def infield(name,base,outcount) :  
  
    playerInfieldFONum = int(hit1.loc[hit1["이름"] == name]['내야뜬공_개수'].values[0])  
    playerHitInp = float(hit1.loc[hit1["이름"] == name]['InP%'].values[0])  
    playerPA = int(hit1.loc[hit1["이름"] == name]['타석'].values[0])  
    playerInfieldRatio = float(hit1.loc[hit1["이름"] == name]['방향_내야%'].values[0])  
    playerHitNum = int(playerHitInp * playerPA / 100)  
    playerInfieldNum = int(playerInfieldRatio * playerHitNum / 100)  
    playerInfieldGround = playerInfieldNum - playerInfieldFONum  
    playerInfieldFORatio = playerInfieldFONum/playerInfieldNum  
  
    playerInfieldFORatio *= 1000  
  
    selector = r.randrange(1,1001)  
  
    if(selector <= playerInfieldFORatio):  
        return 'F0', base,outcount  
    else:  
        return groundBall(name,base,outcount,playerInfieldGround)
```

# Groundball:



```
def groundBall(name,base,outcount,playerInfieldGround):  
  
    playerInfieldHitNum = int(hit1.loc[hit1["이름"] == name]['내야안타_개수'].values[0])  
    playerInfieldHitRatio = playerInfieldHitNum/playerInfieldGround  
    playerInfieldHitRatio *= 1000  
    playerForceOutNum = int(hit1.loc[hit1["이름"] == name]['포스아웃'].values[0])  
-----  
    playerOutPlusRatio = float(team2.loc[team2["이름"] == name]['0+_%'].values[0])  
    playerOutPlusRatio *= 10  
  
    playerDORatio = float(team1.loc[team1["이름"] == name]['병살타%'].values[0])  
    playerDORatio *= 10
```

## [Used data/indicators]

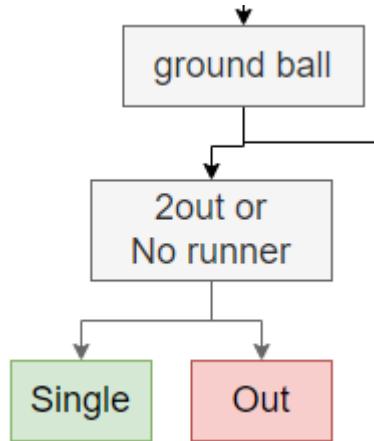
#Infield hits, #On-base due to force-out, Advance rate on outs%, Double play rate in situation%

## [Calculate Infield hits probability]

Infield hits probability% = #Infield hits / #Infield groundball

# Groundball: 2out or no runner

No situations such as double out and productive out(batter is out, but runner advance the base)



```
if outcount == 2 or base == 0:  
    selector = r.randrange(1,1001)  
    if(selector <= int(playerInfieldHitRatio)):  
        base = base<<1  
        base += 1  
        return 'infieldHit',base,outcount  
    else :  
        outcount += 1  
        return 'GO', base, outcount
```

## [Function]

For infield hits probability%, runner and batter advance 1 base  
Else, groundball out

# Groundball: else (double out, Productive out possible)

```
elif base & 1 == 0:  
    selector = r.randrange(1,1001)  
    if(selector <= int(playerInfieldHitRatio)):  
        base = base<<1  
        base += 1  
        return 'infieldHit', base, outcount  
    else:  
        selector2 = r.randrange(1,1001)  
        if(selector2 <= int(playerOutPlusRatio)):  
            base = base << 1  
            outcount += 1  
            return 'O+', base, outcount  
        else:  
            outcount += 1  
            return 'GO', base, outcount
```

```
else:  
    selector = r.randrange(1,1001)  
    if(selector <= int(playerInfieldHitRatio)):  
        base = base<<1  
        base += 1  
        return 'infieldHit', base, outcount  
    elif(selector <= int(playerInfieldHitRatio + playerDORatio)):  
        base -= 1  
        base = base<<1  
        outcount += 2  
        return 'DoubleOut', base, outcount  
    else:  
        selector2 = r.randrange(1,1001)  
        if(selector2 <= int(playerOutPlusRatio)):  
            base = base << 1  
            outcount += 1  
            return 'O+', base, outcount  
        else:  
            outcount += 1  
            return 'GO', base, outcount
```

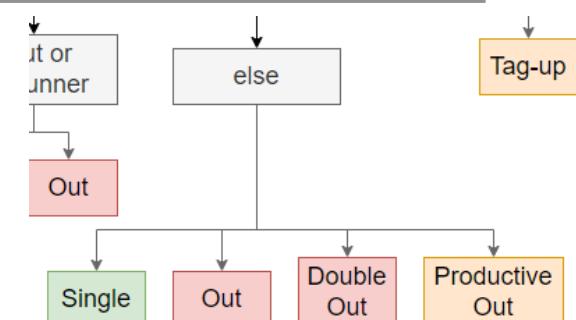
[Function - No 1B runner]: No double play

For Infield hits probability%, runner and batter advance 1 base

Else,

For O+%, runner advance 1 base and batter dies. (Productive Out)

Else, groundball out



# Groundball: else (double out, Productive out possible)

```
elif base & 1 == 0:  
    selector = r.randrange(1,1001)  
    if(selector <= int(playerInfieldHitRatio)):  
        base = base<<1  
        base += 1  
        return 'infieldHit', base, outcount  
    else:  
        selector2 = r.randrange(1,1001)  
        if(selector2 <= int(playerOutPlusRatio)):  
            base = base << 1  
            outcount += 1  
            return '0+', base, outcount  
        else:  
            outcount += 1  
            return 'G0', base, outcount
```

```
else:  
    selector = r.randrange(1,1001)  
    if(selector <= int(playerInfieldHitRatio)):  
        base = base<<1  
        base += 1  
        return 'infieldHit', base, outcount  
    elif(selector <= int(playerInfieldHitRatio + playerDORatio)):  
        base -= 1  
        base = base<<1  
        outcount += 2  
        return 'DoubleOut', base, outcount  
    else:  
        selector2 = r.randrange(1,1001)  
        if(selector2 <= int(playerOutPlusRatio)):  
            base = base << 1  
            outcount += 1  
            return '0+', base, outcount  
        else:  
            outcount += 1  
            return 'G0', base, outcount
```

## [Function - Runner on 1B]

For Infield hits probability%, runner and batter advance 1 base.

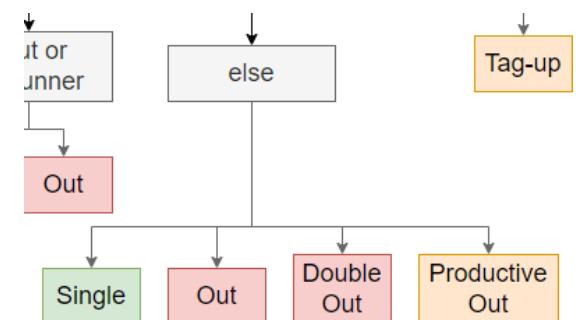
For Double play rate%, batter and 1B runner die.

Else,

For O+%-> runner advance 1 base and batter dies

Else, groundball out

[Limitation] More than 2 runners, 1B runner is fixed to die in double play.



# Limitations of game function

**Cannot consider all of situations.**

**We eliminate some rules..**

- Base steal, Caught stealing, Outs on base
- Balk(Advance the base by pitcher's fault)
- Strike not out, Passed ball, Wild pitch
- Error(Defender's)

THANK YOU  
Q & A